J. Environ. Nanotechnol.
Volume 6, No.1(2017) pp. 75-78
ISSN (Print): 2279-0748
ISSN (Online): 2319-5541
doi:10.13074/jent.2017.03.171235

An Overview of Statistical Pattern Recognition - A Review

C. Kalaiselvi[1], G. Arunasenbagam[2*]

[1]Department of Computer Applications, Tirupur Kumaran College for Women, Tirupur,TN, India.
[*2]Department of Computer Science, Tirupur Kumaran College for Women, Tirupur, TN, India.

Abstract

Pattern recognition is a branch of machine learning that deals with the recognition of patterns and regularities in data, in some cases considered to be nearly synonymous with machine learning.Pattern recognition systems are in many cases trained from labeled as a training data supervised learning but when no labeled data are accessible. In other algorithms can be used to discover previously unknown patterns unsupervised learning. The term pattern recognition is used in data mining and knowledge discovery in databases(KDD) are hard to separate they are largely overlap in their scope. Machine learning is the common term for supervised learning methods and ordinates from artificial intelligence, whereas KDD and data mining have a larger focus on unsupervised methods and stronger connection for using business. Pattern recognition has its origins in engineering, and the term is popular in the context for computer vision conference is named Conference on Computer Vision and Pattern Recognition. In pattern recognition, there may be a higher interest to formalize, explain and visualize the pattern and machine learning are focuses on maximizing the recognition rates. These domains have evolved substantially from their roots in artificial intelligence engineering and statistics and they become increasingly similar by accommodate development and ideas from each other. In machine learning, pattern recognition is the assignment of a label to a given input value. Discriminate analysis was introduced for this same purpose in statistics. Pattern recognition is classification, which attempts to assign each input value to one of a given set of classes. However, pattern recognition is a more general problem that envelope other types of output as well. In sequence labeling it assigns a class to each member of a sequence of values and parsing that assigns a parse tree to an input sentence they describing the syntactic structure of the sentence.

Keywords: *Data clustering; Fuzzy sets; Knowledge discovery in database; Structural; Static.*

1. INTRODUCTION

Statistical pattern recognition relates to the use of statistical techniques for analysis data measurements in order to extract information and make justified decisions. It is a very effective area of study and research, in recent years which has seen many advances. Application such as data mining, web searching, multimedia data retrieval, face recognition, and cursive handwriting recognition, all require robust and efficient pattern recognition techniques. Pattern recognition are very effective field of research intimately bound to machine learning.pattern recognition that desires at building a classify that can regulate the class of an input pattern. This procedure, known as training, corresponds to learning an unknown decision function based only on a set of input-output pairsform the training data or training set in real world applications such as character recognition, a certain amount of information on the problem is usually known advanced. The incorporation of this prior knowledge into the training is the key element that will allow an increase of performance in many applications.

The essential goal of pattern recognition is supervised or unsupervised classification. Among the various structure in which pattern recognition has been traditionally formulated, the statistical approach has been most intensively studied and used in practice. Neural network techniques and methods imported from statistical learning theory have been receiving increasing attention. The design of a recognition system that requires careful attention to the following issues in definition of pattern classes, sensing environment, pattern representation, feature extraction and selection, cluster analysis, classifier design and learning, selection

*** G. Arunasenbagam**
email: Arunagopal2208@gmail.com

of training and test samples, and performance evaluation.In this field, the general problem of recognizing complex patterns with arbitrary orientation, location, and scale remains unsolved. New and emerging applications, such as data mining, web searching, retrieval of multimedia data, face recognition, and cursive script recognition, require robust and efficient pattern recognition techniques. The objective of this review paper is to summarize and compare some of the well-known methods used in various stages of a pattern recognition system and analyse research topics and applications which exciting and challenging field.Statistical pattern theory and techniques, with material drawn from a wide range of fields, including the areas of engineering, statistics, computer science and the social sciences. The book has been updated to cover new methods and applications, and includes a wide range of techniques such as Bayesian methods, neural networks, support vector machines, feature selection and feature reduction techniques.Technical descriptions and motivations are provided, and the techniques.A pattern recognition system based on any PR method mainly includes three mutual-associate and differentiated processes. One is data building; the other two are pattern analysis and pattern classification .Data building convert original information into vector which can be dealt with by computer. Pattern analysis' task is to process the data such as feature selection, feature extraction, data-imension compress pattern classification is to utilize the information acquired from pattern analysis to discipline the computer in order to accomplish the classification.

2. STATISTICALPATTERNRECOGNITION

Statistical pattern recognition has been used successfully todesign a number of commercial recognition systems. In statistical pattern recognition, a pattern is represented by aset of features, or attributes, viewed as a d-dimensional feature vector. Well-known concepts from statisticaldecision theory are utilized to establish decision boundariesbetween pattern classes. The recognition system is operated in two modes: training (learning) and classification The role of the preprocessing module is tosegment the pattern of interest from the background,remove noise, normalize the pattern, and any otheroperation which will contribute in defining a compacter presentation of the pattern. In the training mode, thefeature extraction/selection module finds the appropriatefeatures for representing the input patterns and theclassifier is trained to partition the feature space. Thefeedback path allows a designer to optimize the reprocessing and feature extraction/selection strategies. In theclassification mode, the trained classifier assigns the inputpattern to one of the pattern classes under consideration based on the measured features.Provides a self-contained introduction to statistical pattern recognition.Includes new material presenting the analysis of complex networks.Introduces readers to methods for Bayesian density estimation.Presents descriptions of new applications in biometrics, security, finance and condition monitoring.

Provides descriptions and guidance for implementing techniques, which will be invaluable to software engineers and developers seeking to develop real applications.It describes mathematically the range of statistical pattern recognition techniques.

3. THE RESEARCH OF PATTERN RECOGNITION METHODS

Pattern recognition undergoes an important developing for many years. Pattern recognition include a lot of methods which impelling the development of numerous applications in different filed. Statistical decision and estimation theories have been commonly used in PR for a long time. It is a classical method of PR which was found out during a long developing process, it based on the feature vector distributing which getting from probability and statistical model. The statistical model is defined by a family of class-conditional probability density functions Pr(x|c)(Probability of feature vector x given class c. In SPR, we put the features in some optional order, and then we can regard the set of features as a feature vector. Also statistical pattern recognition deals with features only without consider the relations between features.

4.DATA CLUSTERING

Its aim is to find out a few similar clusters in a mass of data which not need any information of the known clusters. It is an unsupervised method. In general, the method of data clustering can be partitioned two classes, one is hierarchical clustering, and the other is partition clustering. Clustering is the process of making a group of abstract that objects into classes of similar objects.A cluster of data objects can be treated as one grouping cluster analysis, it partition the set of data into groups based on the data similarity and then assign the labels to the groups.The main advantage of clustering over classification adaptable to changes and helps single out useful features that distinguish different groups.

Applications of Cluster Analysis

- Cluster analysis is used in many applications such as market research, pattern recognition, data analysis, and image processing.
- Cluster can also helps marketers discover distinct groups in their customer base. And they can characterize their customer groups based on the acquisition patterns.
- In the field of biology, it can be used to derive plant and animal anatomy, categorize genes with similar functionalities and gain insight into structures inherent to populations.
- Clustering also helps in identification of areas of similar land use in an earth observation database. It also helps in the identification of group of houses in a city according to house type, value, and geographic location.
- Clustering also helps in classifying documents on the web for information discovery.
- Clustering is also used in outsider detection applications such as detection of credit card fraud.
- As a data mining function, cluster analysis serves as a tool to gain insight into the distribution of data to detect characteristics of each cluster.

5. THE APPLICATION OF FUZZY SETS

The process of human being is often fuzzy and uncertain, and the languages of human are often fuzzy. we always give complete answers or classification, so theory of fuzzy sets come into being. Fuzzy set describe the extension and intension of a concept effectively. The application of fuzzy sets in pattern recognition started in where the two basic operations –abstraction and generalization were quite much aimed at by bellan.Two principles proposed by and which Bellancan be think as the general role of fuzzy sets in PR.The PR system based on fuzzy sets theory can confidential thinking process of human being widely.

6. NEURAL NETWORKS

Neural network is developing very fast and first neural networks model MP was purposed especially the Hopfield neural networks and famous BP arithmetic came into being after. It is a data clustering method based on distance measurement also this method is model-inattentive. The neural approach applies biological concepts to machines recognize patterns. The outcome of this effort is the invention of artificial neural networks which is set up by the excitation of the physiology knowledge of human brain. Neural networks is composed of a series of different associate unit. In genetic algorithms applied in neural networks is a statistical optimized algorithms proposed by Holland Neur. PR is a very attractive since it requires minimum of a priori knowledge, and with enough layers and neurons, an ANN can create any complex decision region.

7. STRUCTURAL PATTERN RECOGNITION

The concept of structural pattern recognition is the fourth time and structural pattern recognition is not based on a firm theory which relies on segmentation and features extraction. Structural pattern recognition emphases on the description of the structure, namely explain how some simple sub-patterns compose one pattern. There are two main methods in structural pattern recognition, syntax analysis and structure matching. The basis of syntax analysis is the theory of formal language, the basis of structure matching is some special technique of mathematics based on sub-patterns. When consider the relation among each part of the object, the structural pattern recognition is best. It is different from other methods, structural pattern recognition handle with symbol information, and this method can be used in applications with higher level, such as image interpretation. Structural pattern recognition ways associates with statistic classification or neural networksthrough which we can deal with morecomplex problem of pattern recognition, such as recognition of multidimensional objects.

8. SYNTACTIC PATTERN RECOGNITION

This method major emphasizes on the rules of composition. And the attractive aspect of syntactic methods is its suitability for dealing with recursion. When finish customizing a series of rules which can describe the relation among the parts of the object, syntactic pattern recognition which is a special kind of structural pattern recognition can be used.

9.UNSUPERVISED CLASSIFICATION

In many applications of pattern recognition, it is extremelydifficult or expensive, or even impossible, to reliably label atraining sample with its true category. Consider, the application of land-use classification in remote sensing. In order to obtain the ™ground truth∫ information (category for each pixel) in the image, either the specific site associated with the pixel should be visited or its category should be extracted from a Geographical Information System, if one is available. Unsupervised classification refers to situations where the objective is to construct decision boundaries based

on unlabeled training data.Unsupervised classification is also known as data clusteringwhich is a generic label for a variety of procedures designed to find natural groupings, or clusters, in multidimensionaldata, based on measured or perceived similarities amongthe patterns. Category labels and other information about the source of the data influence the interpretation ofthe clustering, not the formation of the clusters.

10. CONCLUSION

In its broadest sense pattern recognition is the heart of all scientific inquiry, including understanding ourselves and the real-world around us. And the developing of pattern recognition is increasing very fast, the related fields and the application of pattern recognition became wider and wider. In this paper we expatiate pattern recognition in the round, include the definition of PR, the methods of PR, the ompositionof PR system, the related fields of PR and the application of pattern recognition. In addition, it is an importanttrend to use pattern recognition on engineering; we should make efforts on this. And pattern recognition scientists should pay attention to new technique of PR, and enlarge the application areas of PR.

REFERENCES

Baldi, P. F. and Hornik, K., Learning in linear neural networks: A survey, *{IEEE}Trans. Neural Netw.*, 6(4), 837-858(2002). doi: 10.1109/72.392248

Fu, K. S. and Booth, T. L., Grammatical inference: Introduction and survey: Part I, *{IEEE}Trans. Pattern Anal. Mach. Intell.*, 8(3), 343-359(1986). doi:10.1109/TPAMI.1986.4767796

Fu, K. S., Learning control systems: review and outlook, *{IEEE}Trans. Pattern Anal. Mach. Intell.*, 8(3), 327-342(1986). doi:ieeecomputersociety.org/10.1109/TPAMI.1986.4767795

Gelfand, S. B. and Delp, E. J., On Tree Structured Classifiers, Artificial Neural Networks and Statistical Pattern Recognition, 1(1), 51 – 70(1991).

Gelfand, S. B., Ravishankar, C. S. and Delp, E. J., An iterative growing and pruning algorithm for classification tree design, *{IEEE}Trans. Pattern Anal. Mach. Intell.*, 13(2), 163-174(1991). doi:10.1109/34.67645

Jain, A. K. and Chandrasekaran, B., Dimensionality and sample size considerations in pattern recognition practice, *Handbook of Statistics*, 2(39), 835-855(1987). doi: 10.1016/s0169-7161(82)02042-2

Jain, A. K., Mao, J. and Mohiuddin, K. M., Artificial neural networks: A tutorial, *Computer*, 29(3), 31-44(1996). doi: 10.1109/2.485891

Quinlan, J. R., Simplifying decision trees, *Int. J. Man - Machine Studies*, 27(3), 221-234(1987). doi:10.1016/s0020-7373(87)80053-6

Raudys, S. J and Jain, A. K., Small sample size effects in statistical pattern recognition: Recommendations for practitioners, *{IEEE}Trans. Pattern Anal. Mach. Intell*, 13(3), 252-264(1991). doi:10.1109/34.75512

Reed, R., Pruning algorithms - A survey, {*IEEE} Trans. Neural Netw.*, 4(5), 740-747(1993). doi: 10.1109/72.248452

Tsoi, A. C. and Back, A. D., Locally recurrent globally feed forward networks - A critical review of architectures, *{IEEE}Trans. Neural Netw.*, 5(2), 229-239(1994). doi: 10.1109/72.279187

Wood, J., Invariant pattern recognition: A review, *Pattern Recog.*, 29(1), 01-17(1996). doi:10.1016/0031-3203(95)00069-0

Yu G. Smetanin, Neural networks as systems for pattern recognition: A review, Pattern Recognition and Image Analysis, 5(2), 254-293(1995).